

# Nuclear Retention of *COL1A1* Messenger RNA Identifies Null Alleles Causing Mild Osteogenesis Imperfecta

Deborah A. Redford-Badwal,<sup>‡</sup> Mary Louise Stover,<sup>\*</sup> Maurizia Valli,<sup>§</sup> Monique B. McKinstry,<sup>\*</sup> and David W. Rowe<sup>\*</sup>

Departments of <sup>\*</sup>Pediatrics, School of Medicine, and <sup>‡</sup>Pediatric Dentistry, School of Dental Medicine, University of Connecticut Health Center, Farmington, Connecticut 06030; and <sup>§</sup>Department of Biochemistry, University of Pavia, 27100 Pavia, Italy

## Abstract

Osteogenesis imperfecta (OI) is a heritable connective tissue disorder characterized by bone fragility. Most cases of severe OI result from mutations in the coding region of the *COL1A1* or *COL1A2* genes yielding an abnormal collagen  $\alpha$  chain. In contrast, many patients with mild OI show evidence of a null allele due to a premature stop mutation in the mutant RNA transcript. We have previously described a null allele arising from a splice donor mutation where the transcript containing the included intron was sequestered in the nucleus. Here we demonstrate that transcripts from null alleles arising from premature stop mutations are also present in the nucleus and absent in the cytoplasm. Using reverse transcriptase-PCR and single-strand conformational polymorphism of *COL1A1* mRNA from patients with mild OI, we describe three patients with distinct null producing mutations identified from the mutant transcript within the nuclear compartment. A fourth patient with a Gly $\rightarrow$ Arg expressed point mutation exhibits the mutant transcript in both compartments. Defining the distribution of allelic variants of *COL1A1* mRNA in the nuclear and cytoplasmic compartments gives further insight into cell biology of OI and provides a strategy for investigating potential causes of a null allele. (*J. Clin. Invest.* 1996. 97:1035–1040.) Key words: collagen diseases • bone • RNA splicing • polymerase chain reaction • polymorphism, single strand conformational

## Introduction

Osteogenesis imperfecta (OI)<sup>1</sup> is a heterogeneous group of dominantly inherited connective tissue disorders resulting in increased frequency of bone fracture (1). Clinical classification of this group of disorders is divided into four broad categories based primarily on the degree of bone fragility (2). In most cases, the molecular basis of OI is a mutation within either the

*COL1A1* or *COL1A2* gene, the two genes encoding type I collagen (3). An extensive literature has documented that many of the severe forms (OI types II, III, and IV) are caused by a glycine (Gly) substitution within the Gly-X-Y amino acid repeat located within the triple helical domain of this protein (4). The mutant collagen  $\alpha$  chains, when incorporated into the collagen triple helix, compromise the integrity of the higher order collagen structure, particularly in bone (5). A rough correlation between the location of the glycine substitution, degree of helix instability, and severity of bone disease has been demonstrated, although exceptions do exist (6). Thus, the mildest form of OI (type I) may arise from the synthesis of an abnormal  $\alpha$  chain, usually the result of a mutation located within the NH<sub>2</sub>-terminal portion of the helix (7). Mild disease, however, can also occur from a functionally null *COL1A1* allele (8–10). In contrast to a structural mutation that disrupts higher order collagen formation, a null allele merely decreases production of normal type I collagen.

The molecular basis of a null *COL1A1* allele is frequently the result of a premature stop codon arising either directly from a point mutation or indirectly from a frameshift mutation (11), or from a mutation causing an abnormality in mRNA splicing (12). Our analysis of RNA containing a stop codon within an included intron of the *COL1A1* gene (12) and a stop codon in the chicken aggrecan gene (13) demonstrated that the mutation-bearing transcript is retained in the nuclear compartment and does not accumulate in the cytoplasm, precluding its translation and, thus, rendering the allele null.

This work was initiated to determine if nuclear sequestration of one allelic transcript of *COL1A1* mRNA is also present in patients with a null mutation that does not result from a defect in RNA splicing. Our approach uses the presence of an electrophoretic variant within PCR-amplified segments of *COL1A1* mRNA extracted from the nuclear and cytoplasmic compartments to distinguish between allelic transcripts and to determine their presence in the two cellular compartments. The approach provides a strategy that can distinguish fundamentally different molecular mechanisms for mild OI. In this paper, three examples are given illustrating the production of allelic transcripts from functionally null alleles. A fourth example resulting from an expressed mutant protein also documents the ability to distinguish between an expressed and null *COL1A1* mutation.

## Methods

**Fibroblast culture growth and harvesting DNA and RNA.** Dermal fibroblasts from clinically diagnosed mild OI donors were isolated and cultured in Dulbecco's minimal essential media containing 5% FCS and 5% horse serum as previously described (8). Cultures were grown to confluency and supplemented with sodium ascorbate (25  $\mu$ g/ml) 24 h before harvesting. Cells were removed from the plate with 0.25% EDTA/trypsin, washed in serum-containing media, and

Address correspondence to D.A. Redford-Badwal, Department of Pediatric Dentistry, MC 1610, School of Dental Medicine, 263 Farmington Ave., Farmington, CT 06030-1610. Phone: 203-679-3399; FAX: 203-679-4078; E-mail: RBadwal@NSO1.UCHC.edu

Received for publication 27 June 1995 and accepted in revised form 27 November 1995.

1. Abbreviations used in this paper: OI, osteogenesis imperfecta; RT, reverse transcriptase; SSCP, single-strand conformational polymorphism.

*J. Clin. Invest.*

© The American Society for Clinical Investigation, Inc.

0021-9738/96/02/1035/06 \$2.00

Volume 97, Number 4, February 1996, 1035–1040

pelleted. Analysis of collagen chain synthesis and total collagen as a percentage of total protein synthesized was performed by standard radiolabeling and gel electrophoresis (8).

Genomic DNA was isolated from fibroblasts by resuspending the cells in 0.1 M NaCl, 10 mM Tris, pH 8.0, 25 mM EDTA, 0.5% SDS, and 100 µg/ml proteinase K for 4 h at 55°C, followed by sequential phenol and chloroform extractions. The aqueous phase was removed and adjusted to 2.5 M ammonium acetate, 66% ethanol at room temperature. The DNA was collected by spooling, washed with 70% ethanol, dried, resuspended in 400 µl TE (10 mM Tris, pH 7.4, 1 mM EDTA), quantitated spectrophotometrically (absorbency measured at 260 nm), and stored at 4°C.

Nuclear and cytoplasmic RNA were obtained by disrupting the cell membrane with reticulocyte swelling buffer–Triton as previously described (14). Sample processing was rapid and done at 4°C. The nuclear fraction was pelleted by centrifugation (CRU-5000 centrifuge; International Equipment Co., Needham, MA) at 3,000 rpm for 4 min. The supernatant containing the cytoplasmic RNA was extracted using SDS–proteinase K as previously described (15). The nuclear pellet was reextracted in reticulocyte swelling buffer–Triton, repelleted, and subjected to acid phenol extraction (15). Total RNA was extracted by the same method. RNA yield was determined by spectrophotometry and stored at –70°C. The ratio of  $\alpha 1(I)/\alpha 2(I)$  mRNA was performed by a dot hybridization protocol (14).

**cDNA synthesis.** cDNA was synthesized using 5 µg of nuclear, cytoplasmic, or total fibroblast RNA as previously described (10). Either a 35-mer cDNA primer directed to exon 48 (10) or a 33-mer primer (5'-AAGCCGAAGCTCCAACCGGGACAGACGAAG-GAC3'), complementary to nucleotides 4479–4512 in exon 52, was used. After incubating the reactions for 90 min at 37°C, 10 µg of carrier RNA was added, and the cDNA was phenol/chloroform extracted and ammonium acetate/ethanol precipitated. The pellet was washed with 70% ethanol, dried, resuspended in 30 µl water, and stored at –20°C.

**Amplification of fragments.** The entire  $\alpha 1(I)$  mRNA coding region was reverse transcriptase (RT)-PCR amplified (16) as 18 overlapping,  $^{32}$ P-labeled fragments of ~ 300 bp in size (17). Some of the oligonucleotide primers had an artificial cloning site added to the 5' end of the sequence. Table I lists the placement of nucleotide primers and the fragments generated that encompassed the *COL1A1* mRNA. The oligonucleotide primers were made using a DNA/RNA synthesizer (model 394; Applied Biosystems, Inc., Foster City, CA), har-

vested according to recommended procedures (18), and stored at –20°C. The quantity of oligonucleotide primer was calculated spectrophotometrically.

The 10× PCR buffer stock (300 mM tricine, pH 8.4, 50 mM  $\beta$ -mercaptoethanol, 0.1% gelatin, 1% polyoxyethylene-9-lauryl ether, and 5–25 mM  $MgCl_2$  [19]) was stored at 4°C. The optimal  $MgCl_2$  concentration and annealing temperature were determined for each set of PCR primers. The PCR reaction consisted of 1 µl cDNA, 0.25 µg of the 5' forward and 3' reverse primers, 2 µl of the appropriate 10× PCR buffer, 2.5 U Taq polymerase (GIBCO BRL, Gaithersburg, MD), and 1 µCi [ $\alpha$ - $^{32}$ P]dCTP (3,000 Ci/mmol; Amersham Corp., Arlington Heights, IL). This final volume of 16.8 µl was heated to 94°C for 1 min before the addition of 3.2 µl of 0.4 mM dNTPs (20). The reactions were subjected to 30 cycles of 94°C for 30 s, 42–72°C for 15 s, and 72°C for 30 s.

**Single-strand conformational polymorphism (SSCP).** A 0.4-mm-thick, 19.5 × 36.5 cm SSCP gel (8% acrylamide, 10% glycerol, 1.78 M Tris, pH 8.3, 1.78 M boric acid; and 40 mM EDTA [21]) was prerun at 600 V and room temperature for 30 min in 0.5× TBE (44.5 mM Tris, pH 8.3, 44.5 mM boric acid, and 1 mM EDTA). A 2-µl aliquot of the  $^{32}$ P-labeled PCR reaction was added to 8 µl of denaturing loading buffer (95% formamide, 20 mM EDTA, 0.05% xylene cyanol, and 0.05% bromophenol blue). The samples were heated to 94°C for 10 min, and a 3-µl aliquot was loaded into the 5-mm well. A second aliquot was loaded without prior denaturation. The electrophoresis was run at 600 V (constant voltage) for 15 h at room temperature. The gel was dried and exposed to x-ray film (XAR 5; Eastman Kodak Co., Rochester, NY) for 1–4 h at –70°C to visualize the banding patterns from the primary PCR.

**Identification of SSCP bands.** To isolate a SSCP band, a 2-µl aliquot of the original PCR reaction was rerun on another SSCP gel under the same conditions. The wet gel was covered in Saran wrap (Dow Chemical Co., Indianapolis, IN) and placed against x-ray film at room temperature for 1–3 h. The developed film was placed over the gel, and the area of interest was excised. The gel slice was suspended in 50 µl of water, mechanically crushed for 30 s, placed on ice for 15 min, and centrifuged to separate the liquid from the solid phase.

PCR reamplification of the isolated SSCP fragment was performed in a reaction volume of 100 µl, containing 10 µl of 10× PCR buffer and 2 µl of the template. The molar concentration of oligonucleotide primers and nonradioactive dNTPs as well as the reaction conditions were the same as the original PCR. Generation of the expected size product was confirmed by running a 5-µl aliquot in a 6% TBE acrylamide gel followed by ethidium bromide staining. The remainder (95 µl) of the PCR reaction was incubated for 60 min at 37°C with 50 µg/ml proteinase K (22), followed by phenol/chloroform extraction and ethanol precipitation.

The sample was redissolved in 16 µl water, incubated with the appropriate restriction endonucleases, gel purified, and extracted with a gel purification kit (QIAEX; Qiagen Inc., Chatsworth, CA) according to the manufacturer's instructions. Half of the eluted DNA (10 µl) was ligated to an appropriately linearized and phosphatased pBS-SK (Stratagene, La Jolla, CA) vector (50 ng) or pCRII (Invitrogen Corp., San Diego, CA) and transformed into DH5 $\alpha$  *Escherichia coli* cells (23). Colony identification was performed by restriction enzyme digestion of isolated plasmid DNA or by colony lift (24) with a  $^{32}$ P-labeled probe. Plasmid DNA was isolated by alkaline lysis (25) and resuspended in water. Sequencing of the double-stranded DNA was done by dideoxy sequencing (26) using a T7 Sequenase kit (United States Biochemical Corp., Cleveland, OH).

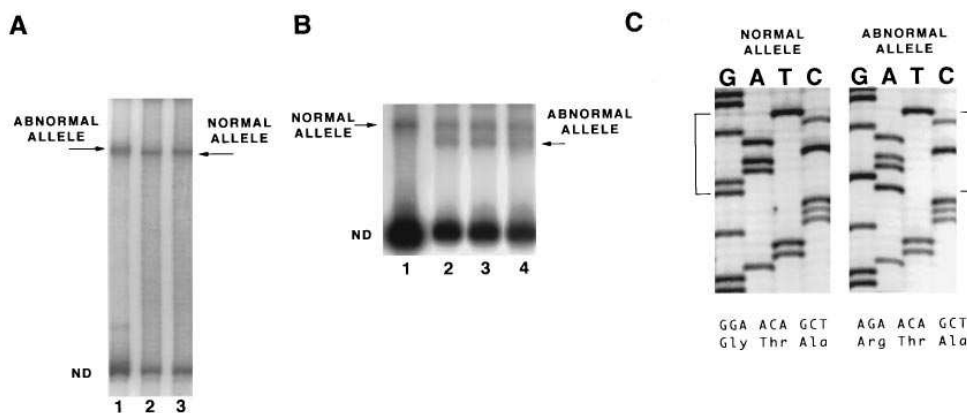
## Results

Our strategy to discriminate between an expressed and a null mutation in mild OI is dependent upon distinguishing the transcripts derived from each *COL1A1* allele. The method uses

Table I. Amplification Strategy

1. Exons 2–6	(306–618 nt)
2. Exons 6–9	(594–808 nt)
3. Exons 9–15	(787–1106 nt)
4. Exons 15–19	(1081–1354 nt)
5. Exons 19–23	(1331–1715 nt)
6. Exons 23–26	(1638–1917 nt)
7. Exons 24–27	(1747–1965 nt)
8. Exons 26–31	(1899–2184 nt)
9. Exons 30–33/34	(2115–2389 nt)
10. Exons 33/34–37	(2368–2629 nt)
11. Exons 36–39	(2542–2767 nt)
12. Exons 38–41	(2689–3000 nt)
13. Exons 41–42	(2978–3219 nt)
14. Exons 43–47	(3195–3516 nt)
15. Exons 47–49	(3492–3782 nt)
16. Exons 49–50	(3781–4128 nt)
17. Exons 50–52	(4119–4377 nt)
18. Exons 51–52	(4360–4502 nt)

Strategy for systematic evaluation of  $\alpha 1(I)$  RNA is listed above. The locations of the oligonucleotides in  $\alpha 1(I)$  mRNA are included.



**Figure 1.** (A) Mild OI nuclear SSCP screen exon 9–15. This figure is from a nuclear screen (exon 9–15) of a number of individuals with mild OI. The differing alleles are indicated by the arrows. Lane 1, patient 196 nuclear cDNA; lane 2, patient control; lane 3, control. Primers used: 9 F 5'CCAGGT-CCCCCTGGAAAGAATG3'; 15 R 5'tcactgatCCAGTAGCA-CCATCATTTCCACGAGC3'. (B) SSCP analysis of cDNA from control (lane 1), patient 196 nuclear RNA (lane 2), cytoplasmic RNA (lane 3), and cDNA from to-

tal RNA of the patient's affected parent (lane 4) is shown here. Note the presence of the abnormal allele in the affected individuals. Additional primers used, 10 F 5'gcaattCCTGGTCGTCCTGGTGAGCGT3'. ND, nondenatured. (C) Genomic sequence, patient 196. Sequencing the antisense strand of the genomic DNA revealed a single base pair change: G to A alters the amino acid Gly to Arg at nt 888.

collagen RNA extracted from the nucleus of cultured fibroblasts as the template to detect an electrophoretic variant that is unique to one patient and absent from other patient and control samples. Once an electrophoretic variant is identified in the nuclear compartment, the cytoplasmic RNA is then examined. The presence of the variant in both compartments suggests an expressed collagen mutation; total RNA preparations may then be used for further analysis. In contrast, presence of the variant in the nucleus alone indicates a null mutation, and further analysis must be restricted to nuclear RNA. In both cases, DNA sequencing will determine whether the unique band represents the causative mutation. An innocuous base change requires further analysis of the patient cDNA derived from the appropriate pool of RNA. We describe the successful application of this method for mutation identification in four patients with OI.

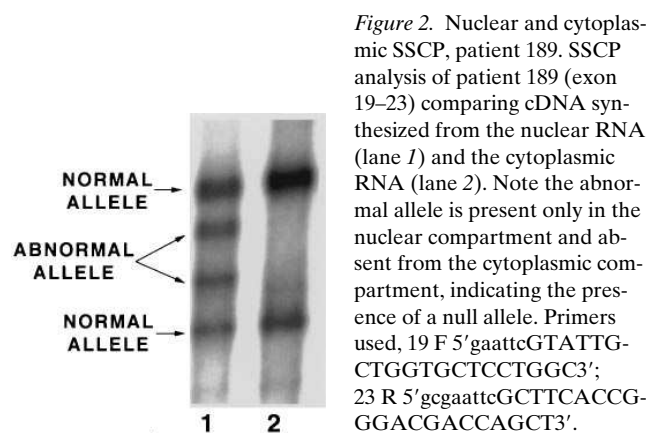
**Patient 196: glycine substitution in exon 11.** The initial SSCP screen of several overlapping *COL1A1* cDNA fragments derived from the nuclear RNA of patients with mild OI revealed a unique band present in the fragment encoded by exon 9–15 of patient 196 (Fig. 1 A). Subsequent SSCP analysis of the same region using the cytoplasmic RNA also demonstrated the presence of the unique band. This evidence of allelism ap-

pearing in both compartments indicated that the mutation leading to the OI phenotype probably resulted from an expressed mutation, and further analysis used total RNA.

Asymmetric restriction enzyme cleavage further localized the mutation to exon 10–15 (data not shown). A subsequent PCR of this region confirmed the original screen and demonstrated a greater band separation (Fig. 1 B). cDNA made from total RNA from the affected parent also confirmed the presence of the two alleles (Fig. 2 B). The identified electrophoretic variant band from patient 196 was excised from the SSCP gel and reamplified using oligonucleotides containing artificially created restriction endonuclease sites. Sequencing of the cloned cDNA as well as genomic DNA revealed a G to A transition at nucleotide 888 (helical codon 79), which changed Gly to Arg in exon 11 in approximately half of the resultant clones (Fig. 1 C).

The synthesis of collagenous proteins by cultured fibroblasts of patient 196 (Table II) revealed low normal levels of total collagen production, whereas the  $\alpha 1(\text{III})/\alpha 1(\text{I})$  chain ratio and the  $\alpha 1(\text{I})/\alpha 2(\text{I})$  mRNA ratio were normal. The migration of the  $\alpha 1(\text{I})$  and  $\alpha 2(\text{I})$  chains on SDS-PAGE were not delayed (data not shown).

**Patient 189: stop codon in exon 19.** The initial *COL1A1* SSCP screen of nuclear RNA from this individual with mild OI revealed two electrophoretic variant bands present in the frag-



**Figure 2.** Nuclear and cytoplasmic SSCP, patient 189. SSCP analysis of patient 189 (exon 19–23) comparing cDNA synthesized from the nuclear RNA (lane 1) and the cytoplasmic RNA (lane 2). Note the abnormal allele is present only in the nuclear compartment and absent from the cytoplasmic compartment, indicating the presence of a null allele. Primers used, 19 F 5'gaattcGTATTG-CTGGTGCTCCTGGC3'; 23 R 5'gcaattcGCTTCACCG-GGACGACCAGCT3'.

**Table II. Protein Examination**

	Collagen synthesis as percentage of total synthesis	$\alpha 1(\text{III})/\alpha 1(\text{I})$ chain	$\alpha 1(\text{I})/\alpha 2(\text{I})$ mRNA
Controls	4–6%	0.14–0.18	1.7–2.3
Patient 196 expressed	$3.5 \pm 0.2\%$	0.16	1.96
Patient 189 null	$2.1 \pm 0.2\%$	1.16	1.4
Patient 263 null	$2.3 \pm 0.3\%$	0.75	0.96
Patient 190 null	ND	0.39	ND

Initial protein analysis of the indicated individuals with osteogenesis imperfecta is listed for comparison. ND, not determined.

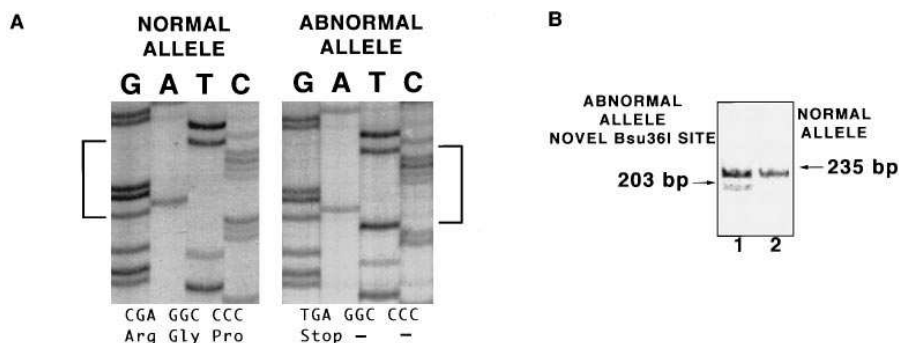


Figure 3. (A) Genomic sequence, patient 189. The sequence of the antisense strand of the confirming genomic sequence is shown here, demonstrating the normal allele (GCA) and the abnormal null allele creating a premature stop codon (TGA) in exon 19, nt 1362. (B) Nuclear and cytoplasmic cDNA-amplified fragments digested with Bsu36I demonstrate the novel restriction site created in the abnormal allele at bp 1363, which removes 32 bp in this allele.

ment encoded by exon 19–23 (Fig. 2). The unique bands were present in the nuclear fraction but absent from the cytoplasmic fraction. Nuclear sequestration of one allele indicates a null mutation. Analysis of collagenous proteins synthesized in the patient's cultured fibroblasts showed that total collagen production was decreased to ~ 50% of control levels (Table II). The ratio of  $\alpha 1(\text{III})/\alpha 1(\text{I})$  collagen was high, suggesting that the reduction in total collagen production was a consequence of a decrease in type I collagen production. Furthermore, the  $\alpha 1(\text{I})/\alpha 2(\text{I})$  mRNA ratio was  $< 2$ , indicative of a decrease in  $\alpha 1(\text{I})$  mRNA.

The electrophoretic variant bands identified in the nuclear RNA were excised from the SSCP gel and reamplified using oligonucleotides containing artificially created restriction endonuclease sites. DNA sequencing of the cloned fragment showed a C to T transition at nucleotide 1362 in exon 19, which changed the amino acid Arg to a stop codon, CGA→TGA (Fig. 3 A). This transition created a new restriction site, Bsu36I, which allowed rapid confirmation of the mutation in the PCR products and demonstrated that the mutation was present in the nuclear RNA but absent from the cytoplasmic RNA (Fig. 3 B). Cloning and sequencing of the genomic DNA isolated from patient 189 confirmed the mutation. The normal and mutant alleles were easily identified by restriction mapping.

**Patient 190: 1-bp deletion in exon 27.** Initial analysis of the  $\alpha 1(\text{III})/\alpha 1(\text{I})$  chain ratio was high, indicative of a null allele

(Table II). The nuclear SSCP screen of the region containing exon 26–31 revealed an additional band present in patient 190 but absent from the other patients (data not shown). The SSCP comparison of nuclear and cytoplasmic RNA from patient 190 revealed the absence of the abnormal band from the cytoplasmic RNA (Fig. 4 A). DNA from the abnormal allele was isolated, reamplified, cloned with the TA cloning kit (Invitrogen Corp.), and sequenced. The sequence of the abnormal cDNA revealed a single base pair deletion, T at nucleotide 1986 in exon 27 (Fig. 4 B). This created a premature stop codon 142 amino acids downstream in exon 33, precluding the exit of this transcript from the nucleus. Confirmation by genomic sequencing was performed (data not shown).

**Patient 263: 2-bp deletion in exon 50.** Patient 263 is a member of a family in which mild OI has previously been described (9). SSCP analysis of nuclear RNA revealed a distinct and intense abnormal band pattern of the exon 49–51 (Fig. 5). The unique bands were limited to the nuclear RNA, whereas the bands derived from the normal allele were present in both compartments. Analysis of collagenous proteins and total collagen mRNA also suggested a null allele. Total collagen production was low, the  $\alpha 1(\text{III})/\alpha 1(\text{I})$  collagen chain ratio was high, and the  $\alpha 1(\text{I})/\alpha 2(\text{I})$  mRNA ratio was low (Table II).

Subsequent cloning and sequencing of abnormal bands revealed a 2-bp deletion in exon 50 (nt 3708–3709), which created a stop codon 21 amino acids downstream in the same exon. This mutation was previously identified by Willing et al. (11) using SSCP analysis of genomic *COL1A1* DNA.

## Discussion

We have demonstrated that allelic transcripts from the *COL1A1* gene containing a premature stop codon are present in the nuclear compartment but absent from the RNA found in the cytoplasm of fibroblasts from patients with mild OI. The

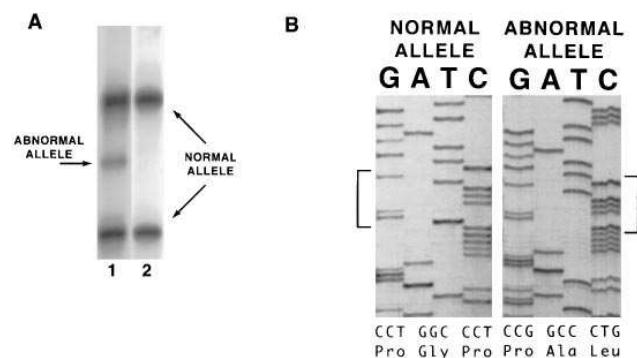


Figure 4. (A) Nuclear and cytoplasmic SSCP, patient 190. SSCP analysis of patient 190 (exon 26–31) comparing cDNA made from nuclear RNA (lane 1) and cytoplasmic RNA (lane 2) demonstrating the presence of an abnormal band sequestered in the nuclear compartment. Primers used, 26 F: AAGGCTGGAGAGCGAGGT; 31 R: TTGCA-CACCACGCTCGCCAG. (B) cDNA sequence, patient 190. Sequencing the antisense strand of the cDNA showed a single base pair deletion of T at nucleotide 1986, which altered the subsequent reading frame, creating a premature stop codon 142 amino acids downstream.

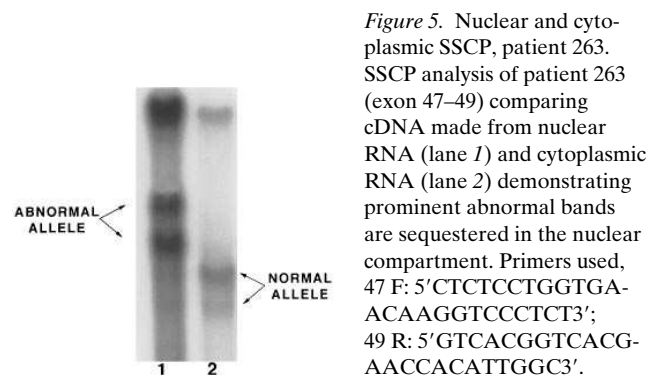


Figure 5. Nuclear and cytoplasmic SSCP, patient 263. SSCP analysis of patient 263 (exon 47–49) comparing cDNA made from nuclear RNA (lane 1) and cytoplasmic RNA (lane 2) demonstrating prominent abnormal bands are sequestered in the nuclear compartment. Primers used, 47 F: 5'CTCTCCTGGTGA-ACAAGGTCCTCT3'; 49 R: 5'GTCACGGTCACG-AACCACATTGGC3'.

literature examining the manner in which cells handle an mRNA containing a premature stop has been inconsistent. Using pulse chase experiments in erythroblasts of individuals with  $\beta^+$ -thalassemia, Benz et al. (27) found nuclear  $\beta/\alpha$  mRNA levels that were almost double those found in nonthalassemic controls, while the cytoplasmic ratio was less than one-half of the controls. This initial evidence for nuclear accumulation of mutant transcripts has been difficult to confirm. Cheng et al. (28) examined human fibroblasts heterozygous for a null mutation of the triosephosphate gene and found decreased levels of triosephosphate mRNA in the nucleus and cytoplasm. They proposed that rapid intranuclear degradation of the mutant transcript occurred at a step after gene transcription and splicing but before nuclear export. Willing et al. were unable to find *COL1A1* transcripts in the nucleus or cytoplasm in fibroblasts of individuals with a functionally null collagen allele using primer extension assays (11). They could not distinguish whether the apparent lack resulted from decreased transcription or rapid degradation. Unlike the above examples in which the RNA with a premature stop is functionally null, Dietz et al. found shortened transcripts from the fibrillin gene in Marfan's syndrome, and in the  $\delta$ -aminotransferase RNA from patients with gyrate atrophy, which resulted from skipping of the exon containing the premature stop codon (29). This occurrence has not been reported for the *COL1A1* gene and would be unlikely to cause mild OI if it arose within the 40 in-frame exons that encode the helical domain (30).

Studies that used transfection to assess RNA processing of a transcript containing a nonsense codon have not always reproduced the naturally occurring mutations and have yielded equally conflicting results. Mutant transcripts that are rapidly degraded in the nuclear compartment (31–34) or transported and degraded in the cytoplasm have been reported (35). With the increasing appreciation of the role the nuclear matrix plays in RNA splicing and subsequent export (36, 37), it cannot be assumed that an RNA from a transfected gene shares the same fate as RNA from an endogenous gene. Analysis of naturally arising null mutations in the *COL1A1* gene in patients with mild OI provides a rich source of material to examine the biological basis for low cellular *COL1A1* mRNA levels.

We have noted variations in the absolute amounts of the electrophoretic variant nuclear transcript as revealed by PCR and SSCP. In our experience, the visual intensity of the novel bands are usually equal to or greater than the normal band, but the ratio may differ between RNA preparations. The variation may reflect the choice of primers used to generate the RT-PCR fragment. Our primers were designed to amplify the entire  $\alpha 1(I)$  mRNA as relatively short but minimally overlapping fragments that may have not been optimized to detect every electrophoretic variant. As newer methods for detection of single base pair mutation in a larger fragment of DNA become available (38), this variability may diminish and the efficiency of our approach may increase. This variability may also represent artifacts related to cell density, or RNA harvesting, or it may reflect true differences in nuclear RNA content. Whether these differences are indicative of a particular kind of mutation is unknown at this time. Greater nuclear accumulation might be expected in null mutations where nuclear export is not detectable. At present, the relative intensity of electrophoretic variants cannot be interpreted except as an indicator of a sequestered nuclear transcript.

To define the molecular basis of a null allele, both protein

and molecular studies are necessary. Protein studies will demonstrate lower levels of type I collagen production when a null allele is present. In cultured fibroblasts from a normal individual, total collagen synthesis accounts for  $\sim 5$ –7% of all protein produced (30). Of the total collagen made,  $\sim 75$ –90% (39) is type I collagen, which results in an  $\alpha 1(III)/\alpha 1(I)$  chain ratio between 0.14 and 0.18. An expressed mutation such as a glycine substitution usually has normal to slightly reduced total collagen synthesis with a normal  $\alpha 1(III)/\alpha 1(I)$  chain ratio. Patient 196, with an expressed mutation, fell within the normal range for the  $\alpha 1(III)/\alpha 1(I)$  chain ratios and showed low normal levels of total collagen synthesis. In contrast, patients 189 and 263, both with null mutations, showed half of the normal total collagen synthesis with an elevated  $\alpha 1(III)/\alpha 1(I)$  chain ratio, indicative of low type I collagen production. Qualitatively, in none of these patients did the mobility of the  $\alpha$  chains deviate from normal.

At the molecular level, a distinction between null and expressed mutations is also apparent. Previously, we have reported that the ratio of  $\alpha 1(I)$  to  $\alpha 2(I)$  mRNA from patients with mild OI was reduced by  $\sim 50\%$  and correlated with low type I collagen synthesis (8). However, the reduction in  $\alpha 1(I)$  mRNA can be variable, probably because of limitations in the hybridization procedures. Willing and co-workers (11, 40, 41), capitalizing on allelic differences in collagen mRNA, show that one allelic transcript was absent from the total RNA in many patients with type I OI. Our paper shows that analysis of nuclear mRNA by RT-PCR followed by SSCP can be used to identify electrophoretic variants derived from the two allelic transcripts. Once identified, it can be determined whether both allelic transcripts are present in the cytoplasmic compartment. Patient 196 demonstrated two allelic transcripts in both compartments and proved to be OI from an expressed mutation. Patients 189, 190, and 263 demonstrated both transcripts in the nucleus but showed only normal transcript in the cytoplasm. The nuclear-retained transcript illustrates the biology of a null mutation and focuses the mutation identification analysis on the mutation-harboring allele.

A null allele may arise from a gene deletion, promoter inactivation, premature stop codons, and disruption of the polyadenylation signal. An additional mutation unique to collagen can occur in the COOH-terminal propeptide by reducing incorporation of the affected  $\alpha 1(I)$  chain into the procollagen molecule (42). By analyzing the quantity of collagen protein and the distribution of the mRNA transcripts from both genes, an expressed allele, a transcribed but null allele, or nontranscribed allele can be distinguished. This information should direct DNA sequencing to the region of the gene most likely to contain the underlying mutation. As this paper illustrates, null alleles that are transcribed but have low type I collagen synthesis show an electrophoretic variant that is sequestered in the nucleus. There are two exceptions to nuclear sequestration that point to unusual causes for a null allele. The first is a null resulting from a lack of transcription of the mutant allele. In this case, low type I collagen synthesis is associated with no evidence of an electrophoretic variant in either the nucleus or cytoplasm in the collagen mRNA, although allelism can be found in the exonic sequences of the genomic DNA. The second is a mutation in the COOH-terminal propeptide affecting chain assembly. In this case, low type I collagen synthesis is associated with an electrophoretic variant in both the nuclear and cytoplasmic compartments.

The pathophysiology that accounts for the low bone mass in OI is still not well understood. Contradictory evidence of a high rate of bone formation/turnover and a low rate of bone formation/turnover have been presented (43, 44, 45). Clinical analysis of patients with mild bone disease resulting from either underexpression of a normal collagen molecule or accumulation of an abnormal collagen molecule may begin to resolve some of these contradictory findings. There are no distinguishing clinical features in terms of fracture frequency, stature, or blue sclera that discriminate patients with an expressed mutation from those with a null mutation. A more detailed analysis of clinical parameters such as bone densitometry, bone histomorphometry, and serum and urine markers for bone turnover, correlated to the molecular basis for defective matrix formation, may provide diagnostic insight for future treatment of this disease.

## Acknowledgments

This work was supported in part by National Institutes of Health grants DE00276 (D.A. Redford-Badwal), AR30426 (D.W. Rowe), and Northeast Osteogenesis Imperfecta Society Grant (D.W. Rowe).

## References

- Byers, P.H. 1993. Osteogenesis imperfecta. Connective Tissue and Its Heritable Disorders Molecular. In Genetic and Medical Aspects. P.M. Royce, and B. Steinmann, editors. Wiley-Liss, New York. 317–350.
- Sillence, D.O., A. Senn, and D.M. Danks. 1979. Genetic heterogeneity in osteogenesis imperfecta. *J. Med. Genet.* 16:101–116.
- Sykes, B., D. Ogilvie, P. Wordsworth, G. Wallis, C. Mathew, P. Beighton, A. Nicholls, F.M. Pope, P. Tsipouras, R. Schwartz, et al. 1990. Consistent linkage of dominantly inherited osteogenesis imperfecta to type I collagen loci: *COL1A1* and *COL1A2*. *Am. J. Hum. Genet.* 46:293–307.
- Byers, P.H., G.A. Wallis, and M.C. Willing. 1991. Osteogenesis imperfecta: translation of mutation to phenotype. *J. Med. Genet.* 28:433–442.
- Kuivaniemi, H., G. Tromp, and D.J. Prockop. 1991. Mutations in collagen genes: causes of rare and some common disease in humans. *FASEB J.* 5: 2052–2060.
- Tsipouras, P. 1993. Osteogenesis imperfecta. In McKusick's Heritable Disorders of Connective Tissue. 5th edition. P. Beighton, editor. Mosby, St. Louis. 281–314.
- Cole, W.G. 1994. Collagen genes: mutations affecting collagen structure and expression. *Prog. Nucleic Acid Res. Mol. Biol.* 47:29–80.
- Rowe, D.W., J.R. Shapiro, M. Poirier, and S. Schlesinger. 1985. Diminished type I collagen synthesis and reduced alpha 1(I) collagen messenger RNA in cultured fibroblasts from patients with dominantly inherited (type I) osteogenesis imperfecta. *J. Clin. Invest.* 76:604–611.
- Barsh, G.S., K.E. David, and P.H. Byers. 1982. Type I osteogenesis imperfecta: a nonfunctional allele for pro $\alpha$ 1(I) chains of type I procollagen. *Proc. Natl. Acad. Sci. USA.* 79:3838–3842.
- Shapiro, J.R., M.L. Stover, V.E. Burn, M.B. McKinstry, A.L. Burshell, S.D. Chipman, and D.W. Rowe. 1992. An osteopenic nonfracture syndrome with features of mild osteogenesis imperfecta associated with the substitution of a cysteine for glycine at triple helix position 43 in the pro $\alpha$ 1(I) chain of type I collagen. *J. Clin. Invest.* 89:567–573.
- Willing, M.C., S.P. Deschenes, D.A. Scott, P.H. Byers, R.L. Slayton, S.H. Pitts, H. Arikat, and E.J. Roberts. 1994. Osteogenesis imperfecta type I: molecular heterogeneity for *COL1A1* null alleles of type I collagen. *Am. J. Hum. Genet.* 55:638–647.
- Stover, M.L., D. Primorac, S.C. Lui, M.B. McKinstry, and D.W. Rowe. 1993. Defective splicing of mRNA from one *COL1A1* allele of type I collagen in nondeforming (type I) osteogenesis imperfecta. *J. Clin. Invest.* 92:1994–2002.
- Primorac, D., M.L. Stover, S.H. Clark, and D.W. Rowe. 1994. Molecular basis of nanomelia, a heritable chondrodystrophy of chicken. *Matrix Biol.* 14: 297–305.
- Genovese, C., and D.W. Rowe. 1987. Analysis of cytoplasmic and nuclear mRNA in fibroblasts from patients with type I osteogenesis imperfecta. *Methods Enzymol.* 145:223–235.
- Chomczynski, P., and N. Sacchi. 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.* 162:156–159.
- Saiki, R.K., D.H. Gelfand, S. Stoffel, S.J. Scharf, R. Higuchi, G.T. Horn, K.B. Mullis, and H.A. Erlich. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science (Wash. DC).* 239:487–491.
- Sheffield, V.C., J.S. Beck, A.E. Kwitek, D.W. Snadstrom, and E.M. Stone. 1993. The sensitivity of single-strand conformation polymorphism analysis for the detection of single base substitutions. *Genomics.* 16:325–332.
- ABI Evaluating and Isolating Oligonucleotides: The Complete Guide (Formerly USER Bulletin 13, 1987). Appendix 5. A13–A14.
- Ponce, M.R., and J.L. Micol. 1992. PCR amplification of long DNA fragments. *Nucleic Acids Res.* 20:623.
- Chou, Q., M. Russel, D.E. Birch, J. Raymond, and W. Bloch. 1992. Prevention of the pre-PCR mis-priming and primer dimerization improves low-copy-number amplifications. *Nucleic Acids Res.* 20:1717–1723.
- Hayashi, K. 1991. PCR-SSCP: a simple and sensitive method for detection of mutations in the genomic DNA. *PCR Meth. & Applic.* 1:34–38.
- Crowe, J.S., H.J. Cooper, M.A. Smith, M.J. Sims, D. Parker, and D. Gewirt. 1991. Improved cloning efficiency of polymerase chain reaction (PCR) products after proteinase K digestion. *Nucleic Acids Res.* 19:184.
- Sambrook, J., E.F. Fritsch, and T. Maniatis. 1989. Molecular Cloning: A Laboratory Manual. 2nd edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY. 1.53–1.89.
- Ausubel, F.M., R. Brent, R.E. Kinston, D.D. Moore, J.G. Seidman, J.A. Smith, and K. Struhl. 1994. Current Protocols in Molecular Biology. Vol. 1. John Wiley & Sons, New York. 6.3.1–6.3.3, Suppl. 24.
- Ausubel, F.M., R. Brent, R.E. Kinston, D.D. Moore, J.G. Seidman, J.A. Smith, and K. Struhl. 1994. Current Protocols in Molecular Biology. Vol. 1. John Wiley & Sons, New York. 1.6.1–1.6.2, Suppl. 24.
- Sanger, F., S. Nicklen, and A.R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA.* 74:5463–5467.
- Benz, E.J., A.L. Scarpa, B.L. Tonkonow, H.A. Pearson, and A.K. Ritchey. 1981. Posttranscriptional defects in  $\beta$ -globin messenger RNA metabolism in  $\beta$ -thalassemia: abnormal accumulation of  $\beta$ -messenger RNA precursor sequences. *J. Clin. Invest.* 68:1529–1538.
- Cheng, J., and L.E. Maquat. 1993. Nonsense codons can reduce the abundance of nuclear mRNA without affecting the abundance of pre-mRNA or the half-life of cytoplasmic mRNA. *Mol. Cell. Biol.* 13:1892–1902.
- Dietz, H.C., D. Valle, C.A. Francomano, R.J. Kendzior, R.E. Pyeritz, and G.R. Cutting. 1993. The skipping of constitutive exons in vivo induced by nonsense mutations. *Science (Wash. DC).* 259:680–683.
- Chipman, S.D., J.R. Shapiro, M.B. McKinstry, M.L. Stover, P. Branson, and D.W. Rowe. 1992. Expression of mutant  $\alpha$ 1(I)-procollagen in osteoblast and fibroblast cultures from a proband with osteogenesis imperfecta type IV. *J. Bone Miner. Res.* 7:793–805.
- Humphries, R.K., T.J. Ley, N.P. Anagnou, A.W. Baur, and A.W. Nienhuis. 1984. Beta O-39 thalassemia gene: a premature termination codon causes beta-mRNA deficiency without affecting cytoplasmic beta-mRNA stability. *Blood.* 64:23–32.
- Takeshita, K., B.G. Forget, A. Scarpa, and E.J. Benz. 1984. Intracellular defect in beta-globin mRNA accumulation due to a premature translation termination codon. *Blood.* 64:13–22.
- Urlaub, G., P.J. Mitchell, C.J. Ciudad, and L.A. Chasin. 1989. Nonsense mutations in the dihydrofolate reductase gene affect RNA processing. *Mol. Cell. Biol.* 9:2868–2880.
- Cheng, J., P.M. Fogel, and L.E. Maquat. 1990. Translation to near the distal end of the penultimate exon is required for normal levels of spliced triosephosphate isomerase mRNA. *Mol. Cell. Biol.* 10:5215–5225.
- Maquat, L.E., A.J. Kinniburgh, E.A. Rachmilewitz, and J. Ross. 1981. Unstable  $\beta$ -globin mRNA in mRNA-deficient  $\beta^0$  thalassemia. *Cell.* 27:543–553.
- Baskin, Y. 1995. Research news: mapping the cell's nucleus. *Science (Wash. DC).* 268:1564–1565.
- Maquat, L.E. 1991. Nuclear mRNA export. *Cell Biol. Int.* 3:1004–1012.
- Youi, R., B.W. Kemper, and R.G. Cotton. 1995. Screening for mutations by enzyme mismatch cleavage with T4 endonuclease VII. *Proc. Natl. Acad. Sci. USA.* 92:87–91.
- Sykes, B., M. Francis, D. Phil, and R. Smith. 1977. Altered relation of two collagen types in osteogenesis imperfecta. *N. Engl. J. Med.* 296:1200–1203.
- Willing, M.C., C.J. Pruchno, M. Atkinson, and P.H. Byers. 1992. Osteogenesis imperfecta type I is commonly due to a *COL1A1* null allele of type I collagen. *Am. J. Hum. Genet.* 51:508–515.
- Willing, M.C., C.J. Pruchno, and P.H. Byers. 1993. Molecular heterogeneity in osteogenesis imperfecta type I. *Am. J. Med. Genet.* 45:223–227.
- Willing, M.C., D.H. Cohn, and P.H. Byers. 1990. Frameshift mutation near the 3' end of the *COL1A1* gene of type I collagen predicts an elongated pro $\alpha$ 1(I) chain and results in osteogenesis imperfecta type I. *J. Clin. Invest.* 85: 282–290.
- Ste-Marie, V.L.G., S.A. Charhon, E.C. Chapuy, and M.C. Meunier. 1984. Iliac bone histomorphometry in adults and children with osteogenesis imperfecta. *J. Clin. Pathol.* 37:1081–1089.
- Baron, R., J.M. Gertner, R. Lang, and A. Vignery. 1983. Increased bone turnover with decreased bone formation by osteoblasts in children with osteogenesis imperfecta tarda. *Pediatr. Res.* 17:204–207.
- Brenner, R.E., U. Vetter, A.M. Bollen, M. Morike, and D.R. Eyre. 1994. Bone resorption assessed by immunoassay of urinary cross-linked collagen peptides in patients with osteogenesis imperfecta. *J. Bone Miner. Res.* 9: 993–997.